

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 6, 2001, 23:27:42 ; Search time 39.77 Seconds
(without alignments)
1088.071 Million cell updates/sec

Title: US-09-494-297-2
File: 3945
Sequence: 1 MKTRFPNKLTNTQRLVLS.....IAGISLGIMGHTIRIKHD 757

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues
Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.0401:*

- 1: /cgnl_8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
- 2: /cgnl_8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
- 3: /cgnl_8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
- 4: /cgnl_8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
- 5: /cgnl_8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
- 6: /cgnl_8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
- 7: /cgnl_8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
- 8: /cgnl_8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
- 9: /cgnl_8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
- 10: /cgnl_8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
- 11: /cgnl_8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
- 12: /cgnl_8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
- 13: /cgnl_8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
- 14: /cgnl_8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
- 15: /cgnl_8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
- 16: /cgnl_8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
- 17: /cgnl_8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
- 18: /cgnl_8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
- 19: /cgnl_8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
- 20: /cgnl_8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
- 21: /cgnl_8/gcgdata/geneseq/geneseqp/AA2001.DAT:*
- 22: /cgnl_8/gcgdata/geneseq/geneseqp/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	181	4.6	905	20	W89421
2	180	4.6	597	21	Y90257
3	178	4.5	898	20	W89413
4	161	4.1	1416	20	Y00211
5	161	4.1	1448	20	Y00210
6	154	3.9	1849	18	W17900
7	152	3.9	1849	19	W56573
8	152	3.9	2516	18	W17899
9	152	3.9	2516	19	W56572
10	149	3.8	1315	20	Y08642
11	149	3.8	2032	20	Y00238

12	149	3.8	2032	20	Y00240	Enterococcus faeca
13	149	3.8	2032	20	Y00242	Enterococcus faeca
14	146	3.7	1185	13	R22675	Collagen binding p
15	145	3.7	1112	20	Y08603	S. pyogenes SF8P-
16	144.5	3.7	1092	19	W41602	Staphylococcus epi
17	142.5	3.6	627	19	W62451	Mycoplasma hyopneu
18	142	3.6	991	21	Y83171	Cell wall protein
19	142	3.6	991	21	Y70120	Staph. epidermidis
20	139	3.5	1231	20	Y00219	Enterococcus faeca
21	139	3.5	1265	20	Y00218	Enterococcus faeca
22	139	3.5	2366	17	R95011	C. difficile toxin
23	139	3.5	2366	17	W68388	Clostridium diffic
24	138	3.5	894	20	W89417	Moraxella catarrha
25	137.5	3.5	2522	20	Y33729	Photothabdus lumn
26	136	3.4	617	21	G30569	Arabidopsis thalia
27	136	3.4	641	21	G30568	Arabidopsis thalia
28	136	3.4	645	21	G30567	Arabidopsis thalia
29	136	3.4	1752	21	G50492	Arabidopsis thalia
30	135	3.4	1228	17	R77673	S-layer protein en
31	135	3.4	1228	18	W22862	Bacillus stearothe
32	135	3.4	1228	20	W93252	B. stearothermophi
33	135	3.4	1228	21	B10625	B. stearothermophi
34	134.5	3.4	2120	21	Y81710	Streptococcus pneu
35	131.5	3.3	921	18	W22863	Bacillus stearothe
36	131.5	3.3	921	21	B10626	B. stearothermophi
37	131	3.3	1116	12	R12083	HMP protein. Baci
38	130.5	3.3	887	21	Y81626	Streptococcus pneu
39	130	3.3	597	18	W20536	H. pylori chaperon
40	130	3.3	638	18	W21012	Bacterial amylase
41	129.5	3.3	1684	12	R14948	Enterococcus pneu
42	128	3.2	621	20	Y00243	Streptococcus pneu
43	128	3.2	666	21	Y81627	Adhesion and penet
44	128	3.2	1394	17	R82768	Pol fragment 2 enc
45	127	3.2	795	21	B03138	

ALIGNMENTS

RESULT 1	
ID W89421	W89421 standard; Protein; 905 AA.
XX	
AC W89421;	
XX	
DT 21-JUN-1999	(first entry)
XX	
DE Moraxella catarrhalis VIM9	lactoferrin binding protein 2 (Lbp2).
XX	
KW Lactoferrin receptor; lactoferrin binding protein; Lbp2;	
KW lbp2 gene; infection; otitis media; sinusitis; conjunctivitis;	
KW pneumonia; bronchitis; tracheitis; emphysema; diagnosis; therapy;	
KW vaccine; Branhamella catarrhalis.	
XX	
OS Moraxella catarrhalis.	
XX	
FH	
FT Key	Location/Qualifiers
FT	Misc-difference 282
FT	/note= "encoded by GAC"
FT	Misc-difference 283
FT	/note= "encoded by TAA"
FT	Misc-difference 284
FT	/note= "encoded by GAC"
FT	Misc-difference 285
FT	/note= "encoded by GGC"
FT	Misc-difference 286
FT	/note= "encoded by AAT"
FT	Misc-difference 287
FT	/note= "encoded by TGG"
FT	Misc-difference 288
FT	/note= "encoded by TTG"
FT	Misc-difference 290
FT	/note= "encoded by CTG"

```

FT Misc-difference 291 /note= "encoded by ACC"
FT Misc-difference 292 /note= "encoded by GAT"
FT Misc-difference 293 /note= "encoded by GAT"
FT Misc-difference 294 /note= "encoded by GTC"
FT Misc-difference 295 /note= "encoded by AAA"
FT Misc-difference 296 /note= "encoded by CGC"
FT Misc-difference 297 /note= "encoded by CCA"
FT Misc-difference 298 /note= "encoded by TTT"
FT Misc-difference 299 /note= "encoded by GAT"
FT Region 435..441 /note= "conserved epitope"
FT W09855606-A2.
XX 10-DEC-1998.
XX 02-JUN-1998; 98WO-CA00544.
XX 08-MAY-1998; 98US-0074658.
XX 03-JUN-1997; 97US-0867941.
XX (CONN-) CONNAUGHT LAB LTD.
XX Du R, Klein MH, Loosmore SM, Wang Q, Yang Y;
XX MPI: 1999-070266/06.
XX N-PSDB; V82021.
XX
XX Lactoferrin receptor genes from Moraxella, especially M. catarrhalis
XX - useful to diagnose Moraxella infection e.g. to detect otitis media
XX due to M. catarrhalis infection and to immunise against such
XX infections
XX
XX Claim 8; Fig 16; 202pp; English.
XX
XX This protein comprises lactoferrin binding protein 2 (Lbp2) of
XX of Moraxella catarrhalis (Branhamella catarrhalis) WH19. It is
XX encoded by the lbpB gene of the lactoferrin receptor (lfr) locus
XX (see V72021) identified in the M. catarrhalis WH19 genome.
XX Immunogenic compositions, including vaccines, based upon expressed
XX recombinant Lbp1 and/or Lbp2 and/or ORF3 proteins (see W89413-21),
XX portions of these, or their analogues, can be prepared for
XX prevention of diseases caused by Moraxella. M. catarrhalis is a
XX causative agent of otitis media and has been associated with
XX sinusitis, conjunctivitis and inflammatory diseases of the lower
XX respiratory tract, such as pneumonia, chronic bronchitis,
XX tracheitis and emphysema.
XX
XX Sequence 905 AA:
SO

```

```

Query Match 4.6%; Score 181; DB 20; Length 905;
Best Local Similarity 20.7%; Pred. No. 9.3e-05;
Matches 165; Conservative 103; Mismatches 262; Indels 266; Gaps 39;

```

```

OY 84 YKQFVADHDLNREGSGSYQYCCFNLKA-----FPLGSDSVK-----WYKHDGI 133
DB 165 ytdkfpkisd1-hlense---hvfdaKamnikiygyaalspaknptymyngqeqn1 218
OY 134 STR--FEDYA-----MSPRIT-----GDELNOKLRAVYNG-----HPQANGIMEGI 174
DB 219 kkkpgdddygnirfygmrelidlnkkgsdtdskrralfitftptlfyngenahtlp-- 276
OY 175 EPLNAIRVTOEAVWYSDNAPISNDE-----SFKRESENLVSTQSLSL----- 219
DB 277 ---kagktfdckvgyfnstrksnegktdckvgyfnstrksnegdlvsaahtlynsfk 333

```

```

OY 220 -----MRQALKQLIDPNLATKMPKQVPDFF-----QLSIFESEDK-- 254
DB 334 ykhtpatsvdtidntlgklsydnpnkqtadgryttsqfctdkkveadryetdakin 393
OY 255 GKYNKGYONLISGL-----VPTKPPYGGDPMPMPNPQOTTSLVLRKAIDYSKL 306
DB 394 gnrtfgtksliddntlnatpfvkelfskkanpnp--dnp-----sdt 434
OY 307 LEGATLQLTGDNVNSFOARVFSSNDI-----GRIEISDGTLYLTENSPAGSIAP 359
DB 435 leggfysesgdL-----agkflsmdnatlvfygkrdkltepvaktyfys-tgle--kp 487
OY 360 IFFKEAGKVYTIIDGK-----QIENPNKEIVPYSVEAYNDE-----EEFSVLTT 405
DB 488 stsfgvneelgldgkrlndevnngidetvrsnkeyeynygrpnkqfckklnasvq 547
OY 406 QNYA-----KRYV-----AK--NKGSSQVYVCFNADLKSPPDSLDGKTIPTDFT 449
DB 548 knpayfgqhdkfyngnyydlasakanklgyvsqdslnksllakypda---kvsldkv 603
OY 450 T-----GEVKYTHIAGRDLFKYTVKPRDTPDPTFK----- 480
DB 604 tkivlqgakdkpytalhaksydhisfgevlyndknpntrsyfvgggqadavstqlpsagk 663
OY 481 -----HIKKVIEKGYREKGAIEVSGLTETQLRAATQAIYFT----- 519
DB 664 ftynglwagyltqkdkyskdedtlkgqikd-----yltkdfipgddddd 712
OY 520 DSAELDKDKLKD-YHGFQDMNDSTLAVAKILVEYVADSNPPLTDLDFIPNNKNGOSLI 578
DB 713 dslasdsqddntnpgdddl-----iasdsqdddtgdgdsddl--gdgaddaa 761
OY 579 GQWHPEDLVDIRMEDKKEVIVP---TNLTL-----RKTYWGLAGDRT 620
DB 762 gkyvynagn---lrpefenkylpneptnekftaladgkkaakfdvndfnstlgtkinder 817
OY 621 KDFHEIELKNNKQELLISQTVTKTNELEFKDKGKATINLKHESITLLOLPBGYSYVKE 680
DB 818 gdlvfdl--knkldgtgtfakadpnyreevy-----mggs-----gflynkld 861
OY 681 TDESGYKRVNSOEVA 696
DB 862 idvkgqfigtungea1 877

```

RESULT 2

```

ID Y90257 standard; Protein: 597 AA.
XX Y90257:
AC Y90257:
XX 19-SEP-2000 (first entry)
DT 19-SEP-2000 (first entry)
XX
XX streptococcus equi fibronectin binding protein, FN2.
DE streptococcus equi fibronectin binding protein, FN2.
KW Fibronectin binding protein; SFS; vaccine; horse; strangles; therapy;
KW equine upper respiratory tract disease; S. equi infection; FN2.
XX
XX Streptococcus equi.
OS Streptococcus equi.
XX WO200037496-A1.
PN 29-JUN-2000.
PD 29-JUN-2000.
XX 21-DEC-1999; 99WO-SE02448.
XX 22-DEC-1998; 98SE-0004491.
PA (GUSG/) GUSG B.
PA (LIND/) LINDARK H.
PA (JACO/) JACOBSSON K.
PA (FRYK/) FRYKBERG L.

```

XX	PI	Gus B, Lindmark H, Jacobsson K, Frydberg L:	
XX	DR	WPI: 2000-442641/38.	
XX	DR	N-PSDB; A30874.	
XX	PT	New protein useful for preparation of vaccines for treatment of	
XX	PT	strangles caused by Streptococcus equi infection, is able to bind to	
XX	PT	mammalian fibrinectin -	
XX	PS	Disclosure: Page 17a-17b; 34pp; English.	
XX	CC	This sequence represents the Streptococcus equi fibrinectin binding	
XX	CC	protein, FN2. The FN2 sequence was used to isolate the S. equi	
XX	CC	fibrinectin binding protein of the invention, designated SFS. SFS binds	
XX	CC	specifically to mammalian fibrinectin or its analogues or fragments. The	
XX	CC	protein, its analogues or fragments may be used for the preparation of a	
XX	CC	vaccine that protects horses against strangles (a world-wide distributed	
XX	CC	and serious disease of the equine upper respiratory tract) caused by the	
XX	CC	S. equi infection. The antibody and/or antiserum may also be used for the	
XX	CC	prophylactic or therapeutic treatment of S. equi infection in mammal,	
XX	CC	especially horses. The use of vaccines containing the fibrinectin binding	
XX	CC	protein provides a more effective protection against S. equi infections,	
XX	CC	with fewer side effects.	
XX	Sequence	597 AA:	
XX	Sequence	597 AA:	
XX	Query Match	4.6%; score 180; DB 21: Length 597;	
XX	Best Local Similarity	20.8%; Pred. No. 5,7e-05;	
XX	Matches	149; Conservative 102; Mismatches 265; Indels 202; Gaps 34.	
QY	68	SEYRWYGEYSYVR-GHPYKQFRVAHDLRVNLEGRSRYQVYCNFKKAPLSDSSVKKW 126	
DB	32	aeglyygnvndgrtgsypf--lyvsnaprkeldeyvvvynfkllklypqdweslysnf 89	
QY	127	-----YKHHDSITKFEEDYANSPRTGTELNOKLAWYMGHPQANANGIMEG- 173	
DB	90	ndisrpyndlpnyekkllygdqfqlkyapdykklidiasalavavinsyptnksglstay 149	
QY	174	-LEPLNATIRVIOEAWWYSDNAPINPDESFRSEESNVSRQSLSLMOQALKOLIDPL 232	
DB	150	hlnndsrkrcvclawlyfsd---sltkyldtggynl-----ndmekkaldfllskqe 200	
QY	233	ATKMKQVPDDFQSLSEFSESDGKDYKNGYONLLSGGLVPTKPPRGDPMPPNQPTTS 292	
DB	201	dskl-ksegsnysldiyvyygsgphnmkdynqllysltlpkcplrp----- 245	
QY	293	VLIRRYAIGDYSKLEGAATLQLTGDNVNSFQARVESNDIGRIELSDGTYTLTEINSPA 352	
DB	246	-----qlygfs-----ghnnglsglegsgsgsgqetne--dqkkgll--lgfhg 284	
QY	353	GYSINL-----PIT-FKVEAGKYTTIIDKQIENPKKEIYEPYSEVAYNDFEESVLTQN 407	
DB	285	glsgegrkrdrplpglksgeagapdt-----pdkrnpdl--qgleagn-----spiveqn 330	
QY	408	YAKFYAANKNKNSSOVVYCFNADLSPDESDGCKTMRPDTFTGCKYKTHNG-----R 461	
DB	331	ygs-----tegyhgqslle--etednppglllsgsgvne--thedrphlngisggslag 383	
QY	462	DLFKTYVPR-----DTPDPTFLHKIKKIVIEKGYREKQCAIYESGLTETQLRAATQ 512	
DB	384	esgetlkrpqrqgsgqpyletledd-----qkmsgsgsggltlesentlkpke 429	
QY	513	LAIYFTDSAEIDKOKLKDYGHGFGMNDSLVAKKIIVEYAADSNRPQQLTDLDFIPNNN 572	
DB	430	ymlgsgqgqletleddtqr--gmsgsgsgtle-----sedtkkpev----- 467	
QY	573	KYQSLIGIQWHEDEDLVDDIIRMEKKEVIRPVTHNLTKRYTGLADRDYDFHFEILKNN 632	
DB	468	-----mlgsg-----gqlldfsen--tqsgmsgsgsdtt-----viedek 500	
QY	633	KOELLISQVTKDKNLEFRDGKATINLKHGESLTLLQGL--PEGYSTVLEKTEIDSEGYKAV 690	

Db	501	kseil	-----	igggagqdfsd	-----	ttppmgsqsggttlivdcttkpckp	545
QY	691	NSQEVANATVSKTGTSDTEIAFENNKPEVPTGYDQK	-----	INGYALIVIAISL	743		
Db	546	apaplvn	-----	de	-----	kpnkgthlpqtsdmkqtlstlsigamemllylclsl	589
RESULT	3						
ID	W89413						
XX	W89413	standard; Protein; 898 AA.					
AC	W89413;						
XX	21-JUN-1999	(first entry)					
XX							
DE	Moraxella catarrhalis	lactoferrin binding protein 2 (lbp2).					
XX							
KW	Lactoferrin receptor; lactoferrin binding protein; lbp2;						
KW	lbp2 gene; infection; otitis media; sinusitis; conjunctivitis;						
KW	pneumonia; bronchitis; tracheitis; emphysema; diagnosis; therapy;						
XX	vaccine; Branhamella catarrhalis.						
OS	Moraxella catarrhalis.						
XX							
FH	Key	Location/Qualifiers					
FT	Misc-difference	632					
FT	/note=	"encoded by AAR"					
FT	Region	430..435					
FT	/note=	"conserved epitope"					
XX							
PN	W09855606-A2.						
XX							
PD	10-DEC-1998.						
XX							
PD	02-JUN-1998;	98WO-CA00544.					
XX							
PE	08-MAY-1998;	98US-0074658.					
PR	03-JUN-1997;	97US-0867941.					
XX							
PA	(CONN-) CONNAUGHT LAB LTD.						
XX							
PI	Du R, Klein MH, Loosmore SM, Wang Q, Yang Y;						
XX							
DR	WPI; 1999-070266/06.						
DR	N-PSDB; V82019.						
XX							
PT	Lactoferrin receptor genes from Moraxella, especially M. catarrhalis						
PT	- useful to diagnose Moraxella infection e.g. to detect otitis media						
PT	due to M. catarrhalis infection and to immunise against such						
PT	infections						
XX							
PS	Claim 8; Fig 2; 202pp; English.						
XX							
CC	This protein comprises lactoferrin binding protein 2 (lbp2) of						
CC	of Moraxella catarrhalis (Branhamella catarrhalis) 4223. It is						
CC	encoded by the lbp2 gene of the lactoferrin receptor (lfr) locus						
CC	(see V72019) identified in the M. catarrhalis 4223 genome.						
CC	Immunogenic compositions, including vaccines, based upon expressed						
CC	recombinant lbp1 and/or lbp2 and/or ORF3 proteins (see W89413-21),						
CC	portions of these, or their analogues, can be prepared for						
CC	prevention of diseases caused by Moraxella. M. catarrhalis is a						
CC	causative agent of otitis media and has been associated with						
CC	sinusitis, conjunctivitis and inflammatory diseases of the lower						
CC	respiratory tract, such as pneumonia, chronic bronchitis,						
CC	tracheitis and emphysema.						
XX							
SQ	Sequence	898 AA;					
Query Match		4.5%; Score 178; DB 20; Length 898;					
Best Local Similarity		18.6%; Pred. No. 0.00015;					
Matches 171; Conservative 118; Mismatches 305; Indels 324; Gaps 422							

```

QY 4 TREPNKINTNTQVLSKNSKRFVTVLGVFLMFLAVTSMVGAKTVEGLVESSTPMAN 63
Db 79 tctpndgndlq-----aqkuaaagflvm-----gkirdsprn-d 115
QY PDSSSE--YRWYG--YESYVRGHP-----YKQKRVADHDLVNL 98
Db 116 pdysndlvqwgqklyvgldahrpddgigtgnlrrptlandikplyfnkfpalsdlhds 175
QY 99 EGSRSYOVYFNLKKAFLGSDSSVK--WYKHDGISTFEEDYAMP-----RIT 147
Db 176 exhr-fopkklntklygygnltprskncltphqgdnknknpvpyenirfyldq 234
QY 148 GDELNOK-----LRAVYNGHPNANGIMEGLEPLN----- 178
Db 235 gssltqmadtpndkdtipkmpmllfynq--enassqjpsagkfnylgnwlysdvkkrr 292
QY 179 AIRYTOEAWVYSDNAPISNPDESFKRESNLTSTQSLT----- 219
Db 293 alasaddrvgylnasgkn-----egdvssaahylngfykhpetygvdfdn 343
QY 220 -MROALKQIDPLNLTAKMPKQVDDF-----QLSIFESBDK--GDKYKGTQNLISGG 269
Db 344 sltqklsydpnpgtqgkyksgfdctkknvtdvyqidakngnrfygtakslvnen 403
QY 270 L-----VPTKPPRPDPMPNPNQPTSVLIRKKAIDGYSKLEGATLQNLGDVNS 321
Db 404 tetapfkelfskampnp--npn-----scllegfygesgdel-- 442
QY 322 FOARVFNSSND-----TGERIELSDGTTLTELNSPAGYSIAEPITFRKAGKYTTIDG 375
Db 443 --agkfinsdnasyvfvgykrtdkdvaktvtyfsagfe--kpsstfvdneltgrlins 498
QY 376 KQLENPKKEIVE---PYSEAYNDF-----EEFSVLITQN-----YAKFTYAK 415
Db 499 kkindavneklidngdiprtdsderdefpwgekkaeftkvvssltqavpayfygndktyf-- 556
QY 416 NKGSSQVYVCFNADLKSPDSEGGKMTPTDFTTGEV----- 453
Db 557 --ngnyddlasssvdklapadvkangskiekypnatlndngytaivlgeakdnkpyta 614
QY 454 ---KYTHIA-GRDLF-----KYTVKRPDTPDPTFKHKIKVI----- 486
Db 615 lraksyghisfgeclyndangtrfsvfyvggradstcllprkagkftynglwagylqkk 674
QY 487 EKGVRKGGALERSG-----LTERQLRAATQOLAIYFT-----DSMLDKDKAKDHGF 535
Db 675 dkysnneetlkkkghdyllted-----flpeddddlasdsqddahgd 722
QY 536 GDMNDSTLAVAKILVEYAOSNPQOLTDLDFLIPNNKYSLSLGTOMHPEDLVIDIMED 595
Db 723 ddi-----lasdsqddadgdddsddl--gdqadaaagkyvhagn-----irpef 767
QY 596 KKEIVPV---THNLT-----RKTVTGLADRTKDFHEIELKNNKQEL 637
Db 768 enkyllrpnethetfaldgknkakfadvdtlmsltgklndergdvfdl--ngkldgt 825
QY 638 SQVYKTDKTLNLEFKDKATINLKHGESLTLOGLPREGSYLVKKEPDSGYKYNVSOEVA- 696
Db 826 gftakadvpyreevg---nngqg-----gflnylkdldvkvqgffgtnggeelag 871
QY 697 -----NATVSKTG 704
Db 872 qlgydkygdgindtaekag 889

```

```

RESULT 4
Y00211
ID Y00211 standard; Protein; 1416 AA.
XX AC Y00211;
XX DT 20-APR-1999 (first entry)

```

```

XX DE Enterococcus faecalis antigenic polypeptide fragment EF104.
XX KW Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
XX KW detection; attenuation; antigenic.
XX OS Enterococcus faecalis.
XX PN WO9850554-A2.
XX PD 12-NOV-1998.
XX PF 04-MAY-1998; 98WO-US08959.
XX PR 14-NOV-1997; 97US-0066009.
XX PR 06-MAY-1997; 97US-0044031.
XX PR 16-MAY-1997; 97US-0046555.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Bailey C, Choi GH, Hromockyj A, Kunsch CA;
XX DR WPI; 1999-070095/06.
XX DR N-PSDB; X20201.
XX PT New isolated Enterococcus faecalis polynucleotides - used to develop
XX PT products for the detection of Enterococcus and for use in vaccines
XX PT for prevention or attenuation of Enterococcus infection
XX PS Claim 9; Page 203; 301pp; English.
XX CC The present sequence represents an antigenic polypeptide fragment
XX CC isolated from Enterococcus faecalis. The present invention describes
XX CC genes, proteins and antigenic polypeptides isolated from E. faecalis.
XX CC The proteins can be used in vaccines for preventing or attenuating an
XX CC infection caused by a member of the Enterococcus genus in an animal.
XX CC They can also be used for detecting Enterococcus antibodies in a sample.
XX CC The nucleotide sequences can be used for detecting Enterococcus nucleic
XX CC acids. Products from the present invention can also be used for
XX CC screening compounds to identify agonists and antagonists of E. faecalis
XX CC protein activity.
XX SQ Sequence 1416 AA;

```

```

Query Match 4.1%; Score 161; DB 20; Length 1416;
Best Local Similarity 21.1%; Pred. No. 0.0059;
Matches 159; Conservative 104; Mismatches 326; Indels 166; Gaps 40;

```

```

QY 39 ALVYSWVGAKTVGLVES---STPMALNPDSSEYRWYGESYVRGHPYKQFRVANDLR 95
Db 484 slstprygnpkakqlyvsgdylepisvnp--lnetawngndq-----ngaysst 531
QY 96 --VNEGSRSYQVYCFNLKKAFLP---LGSDDSVKWKYKHKHDSITFEEDYAMPRTGCD 149
Db 532 ttvsvmgskeklqpnleikvkhpnlyslratkelyfyk---lgt---dytvtprsdgs 584
QY 150 ELNOKLRAVYNGHPNANGIMEGLE-----PLNATRVYNGEAWVYSDNAPISNPDESFK 204
Db 585 vlftrptlneqlpifgfnvpslprkxsiyvdtlprlmsaagllprvdtvtl---nsk 641
QY 205 RESESNLTSTQSLMROALKQIDPLNLTAKMPKQVDDPOLISFESEDKGDKYKNGYON 264
Db 642 rgsertlqssknqflvnarndsfdsisvrtkbpaga--dvlfdlydsn--dyvdslypq 697
QY 265 LLSGLVPTKPPGDPMP--PNQPTTSVLI-----RKVAIGDYSKL--LEGATIO 313
Db 698 ywdrgyfdkpmtrpnagpytltfidentysytdfdgktkryll--eykhangyldvrtly 756
QY 314 LTG-----DNVNSFOARVSSNDIGERIELSDGTTLTFLNLSPAGYSIAEPITTFYVAG 367
Db 757 itgtakepgsnmegsasvgn---ealdlslat-----gaanpflknvtkt 801

```

```

QY 368 KYVT-IIDGK--QIENPNKRIEPIYSVEAYNDFEFSVLTQNYAKFYAKNNKNGSSQV 424
DB 802 tvtknldnkthvknptleltpkgtlnaqlidnsltv-----kypeda 846
QY 425 YCFNADLKSPDSEDDGKWTPTFTTGE---VKYTHIAGDLEFKYTVKPRDTPDFLKH 481
DB 847 y-----slektngakvlfkdytltenltieyntvsanaaglytcttdsetlqmsa 899
QY 482 IKKVI-----EKGYREKGOALEYSGLTETQLRAATQALAIYFTD---SAEL 524
DB 900 skktvtapltlklfsegdaegivylatatfytlnvedengaialkvsfelidnvtlhtatef 959
QY 525 DKDKLDYHGFQDM-NDSTLAIVAKILVEFYAODSNRPOLT-----DLDFRIP----- 569
DB 960 tldkqgysfdaimtgdyltlvtlnvpqeysvdee--yltgkaiklvkqdnqkrlpklk 1017
QY 570 NNNKYO-----SLIGTQWHEDELVDIIRMEDKK-EVIVPYNHNLTKRKYVGLAGDRTKDF 623
DB 1018 dhsrlqvkdsitlyvgdswkpee--nfvsatcdktgqdvpef-----klvsgqvdnkhagv 1070
QY 624 HFEIELKNNKQELLSDTVKTDKTNLEFKDGKATINLKHGSELTLOGLPBGYSYLKRETD 683
DB 1071 ypllysgdegeetayltvkvpgsklevkd--tliyv--gdsww-----pe--dnfvsatdk 1120
QY 684 EGYKV---KYNSEVANATVSKTGITSDETLAFEN 715
DB 1121 tggdvpfekiavgvtn--vdkig---dyeielykn 1150

RESULT 5
Y00210
ID Y00210 standard; Protein: 1448 AA.
XX
XX Y00210;
AC
XX
XX 20-APR-1999 (first entry)
XX
XX Enterococcus faecalis protein EF104.
DE
XX Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
KM detection; attenuation; antigenic.
XX
XX Enterococcus faecalis.
OS
XX
XX WO9850554-A2.
PN
XX
XX 12-NOV-1998.
PD
XX
XX
XX 04-MAY-1998; 98WO-US08959.
PF
XX
XX 14-NOV-1997; 97US-0066009.
PR 06-MAY-1997; 97US-0044031.
PR 16-MAY-1997; 97US-0046655.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Bailey C, Choi GH, Hromockyj A, Kunsch CA;
PI
XX
XX WPI: 1999-070095/06.
DR
XX
XX N-PSDB; X20200.
DR
XX
XX New isolated Enterococcus faecalis polynucleotides - used to develop
PT products for the detection of Enterococcus and for use in vaccines
PT for prevention or attenuation of Enterococcus infection
XX
XX
XX Claim 9; Page 201; 301pp; English.
XX
XX The present sequence represents a protein isolated from
CC Enterococcus faecalis. The present invention describes genes, proteins
CC and antigenic polypeptides isolated from E. faecalis. The proteins can
CC be used in vaccines for preventing or attenuating an infection caused
CC by a member of the Enterococcus genus in an animal. They can also be
CC used for detecting Enterococcus antibodies in a sample. The nucleotide

```

```

CC sequences can be used for detecting Enterococcus nucleic acids.
CC Products from the present invention can also be used for screening
CC compounds to identify agonists and antagonists of E. faecalis protein
CC activity.
XX
XX SO Sequence 1448 AA:

Query Match 4.1%; Score 161; DB 20; Length 1448;
Best Local Similarity 21.1%; Pred. No. 0.0061;
Matches 159; Conservative 104; Mismatches 326; Indels 166; Gaps 40;

QY 39 ALVTSWAGTATVGLVES---STPNAINPDSSEFRKYGESVYRGHPYKQORVANDLR 95
DB 511 slstprygrpkaiqlvsdqylepislvpn--lnaetaagwydq-----ngayssr 558
QY 96 --VNLEGSRSYQYVCFMLKAPF---LGSDSVKKKWKHNDISTFEDYAMSPRTGD 149
DB 559 tlvsvngskkprlgnleikvkhprylslratkelyfyyk---lgt---dytlvtrsdgs 611
QY 150 ELNQKLRVAVYNGHPONANGIMEGL---PLNAIRVTQEAAYWYSDNAPISNPDSFK 204
DB 612 vikftprlneiqrigfnyvpslprkdkelpvdclpilmaseagllrpvdtvtl---nsk 668
QY 205 RESESNLVSTISQSLMQLKQLIDPMLATKMPKQVDDPQLSFSEEDKGDYKNGYON 264
DB 669 rgseltrqskngflvnarndsfslsvrcklpaga--dvlfdlydsn--dgvdslypq 724
QY 265 ILSGGLVTRKPTPGDPPMP--PNOPTQTSVLI-----RKVAIGDSKU---LEGATIQ 313
DB 725 ywdrtgqyfdkprmpnspgytlitfdentnsytlfdgktkryll--eykhangwldvptly 763
QY 314 LTG-----DNVNSFOARVESNDIGERIELSDQTYTLTELSNPDAGYSIAEPITFKVEAG 367
DB 784 ltgakepgsnnnegsavaavqn---ealdllsat-----qaanpclkvctk 828
QY 368 KYVT-IIDGK--QIENPNKRIEPIYSVEAYNDFEFSVLTQNYAKFYAKNNKNGSSQV 424
DB 829 tvtknldnkthvknptleltpkgtlnaqlidnsltv-----kypeda 873
QY 425 YCFNADLKSPDSEDDGKWTPTFTTGE---VKYTHIAGDLEFKYTVKPRDTPDFLKH 481
DB 874 y-----slektngakvlfkdytltenltieyntvsanaaglytcttdsetlqmsa 926
QY 482 IKKVI-----EKGYREKGOALEYSGLTETQLRAATQALAIYFTD---SAEL 524
DB 927 skktvtapltlklfsegdaegivylatatfytlnvedengaialkvsfelidnvtlhtatef 966
QY 525 DKDKLDYHGFQDM-NDSTLAIVAKILVEFYAODSNRPOLT-----DLDFRIP----- 569
DB 987 tldkqgysfdaimtgdyltlvtlnvpqeysvdee--yltgkaiklvkqdnqkrlpklk 1044
QY 570 NNNKYO-----SLIGTQWHEDELVDIIRMEDKK-EVIVPYNHNLTKRKYVGLAGDRTKDF 623
DB 1045 dhsrlqvkdsitlyvgdswkpee--nfvsatcdktgqdvpef-----klvsgqvdnkhagv 1097
QY 624 HFEIELKNNKQELLSDTVKTDKTNLEFKDGKATINLKHGSELTLOGLPBGYSYLKRETD 683
DB 1098 ypllysgdegeetayltvkvpgsklevkd--tliyv--gdsww-----pe--dnfvsatdk 1147
QY 684 EGYKV---KYNSEVANATVSKTGITSDETLAFEN 715
DB 1148 tggdvpfekiavgvtn--vdkig---dyeielykn 1177

RESULT 6
W17900
ID W17900 standard; Protein: 1849 AA.
XX
XX W17900;
AC
XX
XX 29-JAN-1998 (first entry)
XX

```

DE	Photornhabdus luminescens insect toxin TcdAl1.
KM	Insecticide; insect; toxin; pest control; biological control;
KM	Photornhabdus luminescens; TcdA; Southern corn rootworm;
KM	Colorado potato beetle; Western corn rootworm; meal worm;
KM	boll weevil; turf grub; Coleoptera; beet armyworm; black cutworm;
KM	cabbage looper; codling moth; corn earworm; European corn borer;
KM	tobacco hornworm; tobacco budworm; lepidoptera; Hymenoptera;
KM	Diptera, Dictyoptera; Acarina; Homoptera.
XX	
OS	Photornhabdus luminescens strain W-14 (ATCC 55397).
XX	
XX	Key
FT	Location/Qualifiers
FT	1..1849
FT	/label= TcdAl1
FT	Peptide
FT	1..112
FT	/note= "S2 N-terminalus (Claim 30)"
FT	Peptide
FT	156..211
FT	/note= "tryptic peptide (Claim 30)"
FT	Peptide
FT	466..475
FT	/note= "tryptic peptide (Claim 30)"
FT	Peptide
FT	993..1004
FT	/note= "isolated N-terminal peptide (Claim 30)"
FT	Peptide
FT	1297..1312
FT	/note= "tryptic peptide (Claim 30)"
FT	Peptide
FT	1390..1409
FT	/note= "tryptic peptide (Claim 30)"
FT	Peptide
FT	1532..1554
FT	/note= "claimed peptide (Claim 30)"
XX	
XX	W09717432-A1.
PN	
XX	15-MAY-1997.
PD	
XX	
XX	06-NOV-1996; 96MO-US18003.
PF	
XX	
XX	28-AUG-1996; 96US-0705484.
PR	
XX	06-NOV-1995; 95US-0007255.
PR	
XX	28-FEB-1996; 96US-0608423.
XX	
PA	(WISC) WISCONSIN ALUMNI RES FOUND.
XX	
PI	Blackburn MB, Bowen DJ, Cliche TR, Ensign JC, Fatig R;
PI	French-Constant RH, Guo L, Hey TD, Merlo DJ, Orr GL;
PI	Petrelli J, Roberts JL, Rocheleau TA, Schoonover S;
PI	Strickland JA;
XX	
DR	WPI: 1997-281022/25.
DR	N-PSDB: T68844.
XX	
FT	Photornhabdus sp. Insecticidal protein toxins and DNA encoding them -
FT	can be genetically engineered into insect larvae food and plants for
PT	insect control
XX	
XX	Claim 34; Page 201-208; 276pp; English.
XX	
CC	This polypeptide comprises the 209.2 kDa TcdAl1 insecticidal toxin
CC	protein of Photornhabdus luminescens W-14. Its sequence was deduced
CC	from a genomic DNA clone (768844) and includes N-terminal and
CC	tryptic peptide sequences obtained from the isolated protein.
CC	TcdAl1 is a proteolytic cleavage product of rcdA (see W17699).
CC	Claimed toxins of P. luminescens (see W17871, W17884-89, W17699-900,
CC	W18301-06) can be produced by recombinant DNA methods and applied
CC	to, or genetically engineered into, insect larvae food and plants
CC	for insect control. The toxins are particularly effective against
CC	Southern corn rootworm, Colorado potato beetle, Western corn
CC	rootworm, meal worm, boll weevil and turf grub (Coleoptera), beet
CC	armyworm, black cutworm, cabbage looper, codling moth, corn earworm,
CC	European corn borer, tobacco hornworm and tobacco budworm
CC	(Lepidoptera), and are also active against insects of the orders
CC	Hymenoptera, Diptera, Dictyoptera, Acarina and Homoptera. (All
XX	claimed).
XX	

Sequence	1849 AA:
Query Match	3.9%, Score 154; DB 18; Length 1849;
Best Local Similarity	18.3%; Pred. No. 0.03;
Matches 148; Conservative 120; Mismatches 287; Indels 254; Gaps 37;	
157	AVMNGHPONNMGIEGEPLEIAIVTQE-AWVYSD-----NAPISNPDESPKRESE 208
Db	796 aqqlnvapqyvaalv-gldyigsmkectpctyawaenaagvltglnsqdentalhaidesr 854
OY	209 SNLVST---SOLSLMRQALQK-----LIDPNLA-----TKMKQVDPDFOL----- 246
Db	855 saaletyirgavakaaalksrddilyqllldnqysaalktrrlieaala-sigllvynral 913
OY	247 -SIFSEDKG-----DKYNGQYNLLSGGLVPRKPRPGPPMPNPQPTTSVLI 295
Db	914 enveenansqvlsirgffldwdkynkrystwagvsqlyvypenyldpctmigtckmmadal 973
OY	296 RKYA-----IDGYSKLLEGATLOLTGDNVNSFOQARV-----SSNDIGE- 334
Db	974 qvsgsqlnadctvedafmsyltsfegvanlkvtsayhndlnndglltyfiglseadagay 1033
OY	335 -----RIELSDGTYTLT-----ELNSPAG--YSIAEPITFKVAKKYTT-IDGQLE 379
Db	1034 ywrsvdhskfndgkfaanaawsewhkldcpinpykstilprvlyk--srllyllwleqkelt 1090
OY	380 NPNKEIVEEYSVEAVNDFE-----EFSVLTQON-----YAKF 411
Db	1091 kqtgnskdgqytctryelkiahrlrvdgctwtlptltfdvnnkiselkleknrapglycag 1150
OY	412 YYAKN-----KNGSSQVVYCFMADLKSPPDSEDKGKTWP----- 446
Db	1151 ygsgeellvmfyngqddtlslsyknaasmqglylf-admas-----kamtpeqsnvyrdn 1201
OY	447 -----DFTTGGEVYTHIAGRDLEKRYVVKRPDRDP 475
Db	1202 syqgfdlnvrrvnrnrrywgwdyylsmvynngdgyvvgwdyylsmvngdipctlnykaassdl 1261
OY	476 DFFELHIKKVIEEGY-----REKQAIEESGLTETOLRAATOLAIYFTDSAELEDKDKLD 531
Db	1262 klylsprkrlinhnygeggkngcrlmnmkygkigdkfivltslgvnpnssnklmfypvyq 1321
OY	532 YHGFEDMNDSTLAIAKILVEYKQDSNPQOLTDLDFPIPNMK-----YOSLIGTQW----- 582
Db	1322 ysg-----ltsqngngrll--fnrdctyp--skvveawipgaksrltngnaaisdddyatdsal 1373
OY	583 -HPEDLVDIRMEDKEVI-----PYTHNLTL-----RKVTYGACGRTKDFHE----- 626
Db	1374 nkpdcllkygflntcdskycatdvsgrveintaspakvqilvaagsgkeqftcdkvsislp 1433
OY	627 -----IELKNNKQELLSQVFKTKDNTLEF-KDGK-----ATINLKHG 662
Db	1434 spsfedamnyqfnaleidsqglfnfnnsasidvtlfaaedqklgyesfslpvltkvsld 1493
OY	663 ESLTLQGLPEGSYLVKKEPDSGKYVVKVNS-----QEVANATYSKSGITDETLAEENNKE 718
Db	1494 naltlhnengaqm---qmwsgytrctrlntllfargqlvarat--cgidtlismetcgnidqe 1546
OY	719 PVPPTGVDPDKINGYLALIVIAISISLGING 747
Db	1547 p-----qlgkgfyatftvipyynlstbng 1568
RESULT . 7	
ID	W56573
AC	W56573 standard; Protein; 1849 AA.
XX	W56573;
XX	07-AUG-1998 (first entry)
XX	Toxin TcdAII, encoded by the tcdA gene from genomic region tcd.


```

FT Peptide 554..563 /note="tryptic peptide (Claim 30)"
FT Peptide 1080..1092 /note="isolated N-terminal peptide (Claim 30)"
FT Peptide 1385..1400 /note="tryptic peptide (Claim 30)"
FT Peptide 1478..1497 /note="tryptic peptide (Claim 30)"
FT Peptide 1620..1642 /note="tryptic peptide (Claim 30)"
FT Peptide 1938..1948 /note="claimed peptide (Claim 30)"
FT Peptide 1938..2516 /note="N-terminal peptide (Claim 30)"
FT Protein /label="TcdA11
FT Peptide 2327..2345 /note="tryptic peptide (Claim 30)"
FT Peptide 2398..2408 /note="tryptic peptide (Claim 30)"
FT Peptide /note="tryptic peptide (Claim 30)"

MO9717432-A1.
XX 15-MAY-1997.
XX 06-NOV-1996; 96WO-US18003.
XX 28-AUG-1996; 96US-0705484.
XX 06-NOV-1995; 95US-0007255.
XX 28-FEB-1996; 96US-0608423.

(WISC) WISCONSIN ALUMNI RES FOUND.
XX Blackburn MB, Bowen DJ, Cliche TA, Ensign JC, Fatig R;
PI Firench-Constant RH, Guo L, Hey TD, Merlo DJ, Orr GL;
PI Petrell J, Roberts JL, Rocheleau TA, Schoonover S;
PI Strickland JA;
XX WPI; 1997-281022/25.
XX N-PSDB; T68843.
XX Photorhabdus sp. insecticidal protein toxins and DNA encoding them -
PT can be genetically engineered into insect larvae food and plants for
PT insect control
XX Claim 34; Page 185-194; 276pp; English.
XX This polypeptide comprises the 282.9 kDa TcdA insecticidal toxin
CC protein of Photorhabdus luminescens W-14. Its sequence was deduced
CC from a genomic DNA clone (T68843) and includes N-terminal and
CC tryptic peptide sequences obtained from the isolated protein. The
CC 282.9 kDa insect toxin is proteolytically processed into TcdA1 and
CC TcdA11 components (see W17900 and W18301). Claimed toxins of P.
CC luminescens (see W17871, W17884-89, W17899-900, W18301-06) can be
CC produced by recombinant DNA methods and applied to, or genetically
CC engineered into, insect larvae food and plants for insect control.
CC The toxins are particularly effective against Southern corn rootworm,
CC Colorado potato beetle, Western corn rootworm, meal worm, boll weevil
CC and turf grub (Coleoptera), beet armyworm, black cutworm, cabbage
CC looper, codling moth, corn earworm, European corn borer, tobacco
CC hornworm and tobacco budworm (Lepidoptera), and are also active
CC against insects of the orders Hymenoptera, Diptera, Dictyoptera,
CC Acarina and Homoptera. (All claimed).
XX
XX Sequence 2516 AA;

```

```

Query Match 3.9%; Score 152; DB 18; Length 2516;
Best Local Similarity 18.4%; Pred. No. 0.07;
Matches 149; Conservative 119; Mismatches 287; Indels 254; Gaps 38;

```

```

QY 157 AAVYNGHPQANANGIMGLFLNIRYTOE-AVAYYSD-----NAPISNPDSFKRRE 208
DB 884 aqqlnvapqvsalv-gldyqsmkelpyagwenaagvlltaglinsqantlhaidesr 942

```

```

QY 209 SNLVST---SOLSLMRQALKQ-----LIDPNLA-----TKMRQVDDDFOL----- 246
DB 943 saalstyyrlryakaaalksrddllqyllldmqvsaaalktrriaeala-sqlglynral 1001
QY 247 -SIFSEEDKG-----DKYKNGYQNLSSGLVPTPTPTPGDPMPMPNPOTTSVL 295
DB 1002 enveeansgvlsrsgffldwkynkrystwagvsqvlvypenyidpblmrgtkmmda11 1061
QY 296 RKYA-----IGDYSKLEGATLQGLGDVNSQOAVF-----SSNDIGE- 334
DB 1062 qsvsgqlnadvdaefmsyilsfegvanlkvisayhdmnngaltlyfiglsetdaey 1121
QY 335 -----RIEISDGTYYLT-----ELNSPAG--YSIAEPIFEKVEAGKYVT-IDKQIE 379
DB 1122 ywrsvshskfndgkfaanaawehkldcplnpykstrlpvlyk--srlyllwleqkelt 1178
QY 380 NPNKEIPEPSVEAYNDFE-----EFSVLTGN-----YAKF 411
DB 1179 kqlgnakdgyqtecdryelkiahirygdtwntplftdvknkkselkleknrapjlycag 1238
QY 412 YYAKN-----KNGSSOVVYCFNADLKSPDSDGKTMTPT----- 446
DB 1239 yggedcllvmfynqgdlslsyknaamgilylf-admas-----kdmpegsnyrdn 1289
QY 447 -----DPTTGEVRYTHIAGRDLPKRYVKPRDTP 475
DB 1290 syqgfetnnvrrvnnryaedyeipssvssrkdgygwdyllsmynngdlptltnkaasdl 1349
QY 476 DPFLEKHIKVIKGY---REKQAIIEYSGLETQLRRAATOLAIFYTSAELDKRKLKD 531
DB 1350 klyispkrlilhnryegqkrngcnlmkykldgkflivylslgvnpnnsnklmfypvyq 1409
QY 532 YHFGGMNSTLAVAKIIEVYAQDSNPQOLTDLDFIIPNNK---YOSLIGNQW----- 582
DB 1410 ysg-----nsglmggrll--fhndtlyp--skwea1pgaksltnqnaa1gdyatda1 1461
QY 583 -HPEDLVDIIRMEKKKEVI-----PYTHNLTLRKT-----VTGLADRT---KDFHFE- 626
DB 1462 nkpdcllkylfmdtsqgtatdvsqpyelntalspakvqllvakvgeqftadkvs1qp 1521
QY 627 -----TELKNNKQELISQVKTDTNLEF-KDGK-----ATYWLKNG 662
DB 1522 spsfdemnyqfnaleldsgqlnfinsasldvftaadykrlgyes1s1p1tkvs1cd 1581
QY 663 ESLLTLOGLEGVSYLVKFNDSGKYKVVNS---QEVANATYSKGTITGDETLAPENKE 718
DB 1582 natllhnenagaym-----qwsytrtrlnllfarglvarat---tgldtlsmenqn1ge 1634
QY 719 PVPVPTGVDDKINGYIALIYIAGISLGIWG 747
DB 1635 p-----qlgkgyatftv1ppynlsthg 1656

RESULT 9
W56572 9
W56572 standard; Protein; 2516 AA.
XX
AC W56572:
XX
XX 07-AUG-1998 (first entry)
XX
XX Toxin TcdA, encoded by the tcdA gene from genomic region tcd.
XX
XX Photorhabdus luminescens W-14; nematode; symbiotic;
XX Heterorhabditis; tca; tcb; tcd; insecticidal activity; toxin;
XX Lepidoptera; Coleoptera; Hymenoptera; Diptera; Dictyoptera; Acarina;
XX Homoptera; Southern; Western corn rootworm; Colorado potato beetle;
XX mealworm; boll weevil; turf grub; beetle armyworm; black cutworm;
XX cabbage looper; codling moth; corn earworm; European corn borer;
XX Tobacco hornworm; budworm.
XX
XX Photorhabdus luminescens.
XX
XX

```


CC Enterococcus faecalis. The present invention describes genes, proteins
 CC and antigenic polypeptides isolated from *E. faecalis*. The proteins can
 CC be used in vaccines for preventing or attenuating an infection caused
 CC by a member of the Enterococcus genus in an animal. They can also be
 CC used for detecting Enterococcus antibodies in a sample. The nucleotide
 CC sequences can be used for detecting Enterococcus nucleic acids.
 CC Products from the present invention can also be used for screening
 CC compounds to identify agonists and antagonists of *E. faecalis* protein
 CC activity.

XX Sequence 2032 AA;

Query Match 3.8%; Score 149; DB 20; Length 2032;
 Best Local Similarity 19.0%; Pred. No. 0.083;
 Matches 178; Conservative 117; Mismatches 330; Indels 310; Gaps 44;

QY 19 LSKNSKREFTVLGVFLMIFALVTSWY-----GAKTVFG-----LVSSRPNAI 62
 DB 1098 lktamdettlllgahfqlwdqaktqylregltvdagvltfgglpggqyllvetkap--- 1154
 QY 63 NPDSSSEYRWYGYESYRGHPYKQFRVAHDLRVNLEGSRSYGVYCFNLKAPPLGSDSS 122
 DB 1155 -----egytsdelakgrvlttdeetsaagap-tlikndvkvflekmdex 1200
 QY 123 VKRW----YKKGIGISIKF---EDYAMSPRITGDELNOKRLAIVYNGHPONANGIME--G 173
 DB 1201 gkhlvnafrfklehavltpfhweevplapdr-----nangqlvnds 1242
 QY 174 LEP-LNAIRVTOEAVWYYSNAP-----ISNPDSFPRSESNN 210
 DB 1243 lkpjlygfteieaptqylltctpkrfivtntsgqldrvhvklnlygssellkkdqagn 1302
 QY 211 LVSTQSLSM--RQALKQ--LIDPNLATKMPKQVDDFOLISIFESDKDKYKNGYQNL 265
 DB 1303 plagaetsvidtggavrenlvsdangkvvtclapqkyf-----vetkap----- 1349
 QY 266 LSGGLVPTKPPPTGDDPPMPNQPOT-----TSVLIRKVALIGYSKILLEGATLQ 313
 DB 1350 -agyllntepafllaasdgkpatvlatanfvyngqlaklikkdvg--hllsgatfk 1405
 QY 314 LITDDNNSFCARFSSNDIGERI--ELSDGTYLTELNSAGYSI-AEPTTFV---EAG 367
 DB 1406 vldakgetlqgl-tlcngeqlvaehlapqkyrfvetkapqyllntprpfelaeknaq 1464
 QY 368 K-----VYTIIDGKOIENPKKEIIEPYSVAYNDFEESVLTQNTYAK-- 410
 DB 1465 kpravvaadnfvykgafqkvtknsadqlagav---fely-dhmkqslgltatgskdg 1519
 QY 411 -----FYAAKKK-----NGSSQVYVC-----FNADLKSPDSDGCKMTPTDF 448
 DB 1520 kllfrdlapqlyykelkapkllpdgadylllypelvkvelfgdgfdgdel-----f 1569
 QY 449 TTGEVKTTHAGDLEFYTKPRDTPDFLKHKKVIEKGYR-----EKGOAT--- 497
 DB 1570 qlg--afanfkgrvaavlikkidanap)pgtlfklyr--lengekiferevtaekdgsame 1625
 QY 498 -----EVSGLTETOLRAATOLAIYV-----TDSAEIDKDKLKDYHG--FG---DMND 540
 DB 1626 dlagsyeldeladatqdylnkqpylfvkvknsndkqpldelefvyqaevmgrkvneqg 1685
 QY 541 STLAIVAKILVEYKODSNPPOLDLDFEIPNNKTYOS----- 576
 DB 1686 qtlagavfaalnadednqpgspitflnragkvseltlctkgelyakglineghyvlvet 1745
 QY 577 -----LIGQWPEDELVDIIRMEDKKEVPVTHNLTFRKTVGGLAGDRKDPHFELK 630
 DB 1746 kaptqylidctllp---fdvtaqlgkeqplaidllinygqla-----qitke 1789
 QY 631 NKKOELLSTV--KTDKTNLEFRDGRATL-NLKAGESLTQLGEPGY-----SYLV 678
 DB 1790 netgealagavfvkldetg-qtvdgqtnlmsdgkvkvlaknlapqkyrfvetqaptsyll 1848

QY 679 KETDSEGYK-----VKVNSQAEVANA----- 698
 DB 1849 netpsasfltakndqgkpatvwlkapfinygsaaklvkldqkmalagaeftvldaetqg 1908
 QY 699 TVSKTGITSDETLAFENKKEPVVPTGVDOKI-NGY 732
 DB 1909 tvasrlsrdnglvyqnvlhlpqkyrfvetkapdgy 1943

RESULT 12

ID Y00240 standard; Protein; 2032 AA.

AC Y00240;
 DT 20-APR-1999 (first entry)
 XX Enterococcus faecalis protein EF124.
 DE Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
 KW Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
 KM detection; attenuation; antigenic.
 XX Enterococcus faecalis.

XX W09850554-A2.

XX 12-NOV-1998.

XX 04-MAY-1998; 98WC-US08959.

XX 14-NOV-1997; 97US-0066009.

XX 06-MAY-1997; 97US-0044031.

XX 16-MAY-1997; 97US-0046655.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Bailey C, Choi GH, Hromocky J A, Kunsch CA;

XX WPI; 1999-070095/06.

XX DR N-PSDB; X20230.

XX New isolated Enterococcus faecalis polynucleotides - used to develop

XX PT products for the detection of Enterococcus and for use in vaccines

XX PT for prevention or attenuation of Enterococcus infection

XX Claim 9; Page 228-229; 301pp; English.

XX The present sequence represents a protein isolated from

XX Enterococcus faecalis. The present invention describes genes, proteins

XX and antigenic polypeptides isolated from *E. faecalis*. The proteins can

XX be used in vaccines for preventing or attenuating an infection caused

XX by a member of the Enterococcus genus in an animal. They can also be

XX used for detecting Enterococcus antibodies in a sample. The nucleotide

XX sequences from the present invention can also be used for screening

XX compounds to identify agonists and antagonists of *E. faecalis* protein

XX activity.

XX Sequence 2032 AA;

Query Match 3.8%; Score 149; DB 20; Length 2032;

Best Local Similarity 19.0%; Pred. No. 0.083;
 Matches 178; Conservative 117; Mismatches 330; Indels 310; Gaps 44;

QY 19 LSKNSKREFTVLGVFLMIFALVTSWY-----GAKTVFG-----LVSSRPNAI 62
 DB 1098 lktamdettlllgahfqlwdqaktqylregltvdagvltfgglpggqyllvetkap--- 1154
 QY 63 NPDSSSEYRWYGYESYRGHPYKQFRVAHDLRVNLEGSRSYGVYCFNLKAPPLGSDSS 122
 DB 1155 -----egytsdelakgrvlttdeetsaagap-tlikndvkvflekmdex 1200

```

QY 123 VKKW----YKKGDISSTKE---EDYAMSPRITGDELNOKLRAVMYNGHPONANGIME--G 173
DB 1201 gkklvnarfklhnavtftthweevlplardt-----nanqglevds 1242
QY 174 LEP-LNAIRVTOEAWYIYSDNAP-----ISNPDEFKRESEN 210
DB 1243 lkpjlyqfteleaptyllidtpkrfivtqntsgqirdvhvkmlynqgsaelikkdaqn 1302
QY 211 LVSTQSLSLM---ROALKO--LIDPLATKMPKQVDDPOLSFEESEDDGDKRNGYQL 265
DB 1303 plagaefsvldtggavrehlvsdangkvvtldlapkygf-----vetkap----- 1349
QY 266 LSGGLVTRKPTPGDPMPNPOT-----TSVLIRKVAIGDYSKLLEGATLQ 313
DB 1350 -agyllntepsaftlaasdrqkpatvlatanfynygtaklikkdng---hllsgatfk 1405
QY 314 LTGDVNVNFOARVFSSNDIGERI--ELSDGYTLTELNSPAGYST-AEPTTRKV---EAG 367
DB 1406 vldakgetlqtgl-ttngngelvaehlapgkyrfvetkaptyllnttpvfeleaknag 1464
QY 368 K-----VYTIIDGKQIENPNKEIPEYSVEAVNDEFEESVLTQONAK-- 410
DB 1465 kpavvasdnfvsygaigvltunsadqplagav---fely-dhmkqslglatatsgkqg 1519
QY 411 -----FYAKNK-----NGSSQVVC-----FNADLKSPDESDGKTMTPDF 448
DB 1520 klffrdlapgtytykkelkpklpdgsdylllypelvkvvelrgdtkgdpel-----f 1569
QY 449 TTGEVKTTHIAGDLEFKYTVKPRDPTDFLKHIVIEKGR-----EKGQAI--- 497
DB 1570 qly--afanfkgvavfkklidananpipygtlflklyr--lengeklifervlaekdgslame 1625
QY 498 -----ESGSLTEPQLRATOLAIYF-----TDSAEIDKDKLKYHG--FG---DMND 540
DB 1626 dlagasyeideidatqgylvnkqplfvyvkknsndkqpldelefvygaewmrkvneg 1685
QY 541 SFLVAKIIVERAODSNPOLTDLPFIIPNNKYYQS----- 576
DB 1686 qtlagvfaalnadeqngqspitflnragekvseittdkgeiyakglineghyvlvet 1745
QY 577 -----LIGTOMHPEDLVDIRMEKKEVIPTVHNLTLRKTVGLAGDRKDHFEIELK 630
DB 1746 kaptgylldtclhr---fdvtaqlgkcpialgldllyngqta-----qitke 1789
QY 631 NNMKQELISQTV--KTQKTMLEFKDGKATI--NLKHGESLTLQGLPEGY-----SYLV 678
DB 1790 netgealagavfkvldetg-qvvdqgtunmsdkqgkviahnlapgtyrfvetqaptsyll 1848
QY 679 KETDSEGYK-----VKVNSQEFVANA----- 698
DB 1849 netpsasfllakngkqkpatvllkapflinygsaaklvkldqgnalagaefkvtldaetg 1908
QY 699 TVSKGITSDELTAEPNNKEPVPTGVDOKI--NGY 732
DB 1909 tvarslrsdngqlyvnhlpqgkyfvetkekapgy 1943

```

RESULT 13

```

ID Y00242 standard; Protein; 2032 AA.
AC Y00242;
XX 20-APR-1999 (first entry)
DE Enterococcus faecalis protein EF125.
XX Enterococcus faecalis: infection; vaccine; immune response; diagnosis;
KM detection; attenuation; antigenic.
OS Enterococcus faecalis.
XX
XX WO9850554-A2.
PN

```

```

XX 12-NOV-1998.
PD
XX 04-MAY-1998; 98MO-US08959.
PF
XX 14-NOV-1997; 97US-0066009.
PR 06-MAY-1997; 97US-0044031.
PR 16-MAY-1997; 97US-0046655.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Bailey C, Choi GH, Hromockyj A, Kunsch CA;
DR WPI: 1999-070095/06.
DR N-PSDB: X20232.
XX
PT New isolated Enterococcus faecalis polynucleotides - used to develop
PT products for the detection of Enterococcus and for use in vaccines
PT for prevention or attenuation of Enterococcus infection
XX
PS Claim 9; Page 232; 301pp; English.
XX
CC The present sequence represents a protein isolated from
CC Enterococcus faecalis. The present invention describes genes, proteins
CC and antigenic polypeptides isolated from E. faecalis. The proteins can
CC be used in vaccines for preventing or attenuating an infection caused
CC by a member of the Enterococcus genus in an animal. They can also be
CC used for detecting Enterococcus antibodies in a sample. The nucleotide
CC sequences can be used for detecting Enterococcus nucleic acids.
CC Products from the present invention can also be used for screening
CC compounds to identify agonists and antagonists of E. faecalis protein
CC activity.
XX
SQ Sequence 2032 AA;

```

Query Match 3.8%; Score 149; DB 20; Length 2032;
 Best Local Similarity 19.0%; Pred. No. 0.083;
 Matches 178; Conservative 117; Mismatches 330; Indels 310; Gaps 44;

```

QY 19 LSKNSKRFVTLVGVLMFALVTSNV-----GAKTVG-----LVESSPNAI 62
DB 1098 lktandettllagahglwdqaktvlgrevtdavgtlftgglpaggylvetkap--- 1154
QY 63 NPDSSSEYRWYGESYVRGRHPYKOFVNAHDLRVNLGSRVQVCFNKKAFPLGSDSS 122
DB 1155 -----egytvdelakgrvltideetsesgqp-tllndvnxflkmdck 1200
QY 123 VKKW----YKKGDISSTKE---EDYAMSPRITGDELNOKLRAVMYNGHPONANGIME--G 173
DB 1201 gkklvnarfklhnavtftthweevlplardt-----nanqglevds 1242
QY 174 LEP-LNAIRVTOEAWYIYSDNAP-----ISNPDEFKRESEN 210
DB 1243 lkpjlyqfteleaptyllidtpkrfivtqntsgqirdvhvkmlynqgsaelikkdaqn 1302
QY 211 LVSTQSLSLM---ROALKO--LIDPLATKMPKQVDDPOLSFEESEDDGDKRNGYQL 265
DB 1303 plagaefsvldtggavrehlvsdangkvvtldlapkygf-----vetkap----- 1349
QY 266 LSGGLVTRKPTPGDPMPNPOT-----TSVLIRKVAIGDYSKLLEGATLQ 313
DB 1350 -agyllntepsaftlaasdrqkpatvlatanfynygtaklikkdng---hllsgatfk 1405
QY 314 LTGDVNVNFOARVFSSNDIGERI--ELSDGYTLTELNSPAGYST-AEPTTRKV---EAG 367
DB 1406 vldakgetlqtgl-ttngngelvaehlapgkyrfvetkaptyllnttpvfeleaknag 1464
QY 368 K-----VYTIIDGKQIENPNKEIPEYSVEAVNDEFEESVLTQONAK-- 410
DB 1465 kpavvasdnfvsygaigvltunsadqplagav---fely-dhmkqslglatatsgkqg 1519
QY 411 -----FYAKNK-----NGSSQVVC-----FNADLKSPDESDGKTMTPDF 448

```

```

Db 1520 kllfrldapgttyyykkelkpkpdpdgadyillypelvkvvelrtdgdkgdpel-----f 1569
QY 449 TUGEVKTYTHIAGDLKFRVYVKKPRDPTFLKHKKVIEKGR-----EKGAQI--- 497
Db 1570 qlg--afanfkgvafekklidanaplpgtfklyr--lengekiferevtaekdgsllame 1625
QY 498 -----EVSGLTERQLKRAITAIYF-----TDSAEIDPKDLKDYHG--FG---DMND 540
Db 1626 dlgaagsyeldeladatgylyvknqpiyfvvkkksndkqpldellefvygaevmrgkvneqg 1685
QY 541 SFLAVAKIIEVEYADSNPPOLDLDFEIPNNNKYQS----- 576
Db 1686 qtlagavfaylnadegnqpgspitflnragekvsetlctkgeiyakqlneghyylvet 1745
QY 577 -----LIGTQWHPEDLVDIRMEDKREVIPTVHTNLTKRTVTGLAGDRTKDFHFEIELK 630
Db 1746 kaptgyllldtllhp---fdvtaqlgkqplalqdllyngqta-----qltke 1789
QY 631 NNKQELLSQTV--KTQKTNLEFKDGKATI-NLKHGESLTLQGLPEGY-----SYLV 678
Db 1790 netgealagavfkvldetg-qlvdgqtnlmsdkqgkvliaknlapgtvrfvetqaptsyll 1848
QY 679 KETDSEGYK-----VKVNSQEVANA----- 698
Db 1849 netpsasftakngkqkpatvvlkapflnygsaaklvkldqgknaalagaefkvtadaetqg 1908
QY 699 TVSKTGITSDETIAFENKPEVPVPTGVDOKI-NGY 732
Db 1909 tvarslrsdngqlygvnhlqpgkytvetkaptgy 1943

```

RESULT 14

```

R22675 ID R22675 standard; Protein; 1185 AA.
XX AC R22675;
XX DT 04-DEC-1992 (first entry)
XX DE Collagen binding protein.
XX KM CBP; collagen binding protein; mastitis; arthritis.
XX OS Staphylococcus aureus.
XX FH Key Location/Qualifiers
FT Region 30..1185
FT Peptide /note="claim 11; page 29-30"
FT 1..29
FT /label= sig_peptide
FT 30..534
FT /label= A
FT /note="see CC"
FT Region 535..721
FT /label= B1
FT /note="see CC"
FT Region 722..908
FT /label= B2
FT /note="see CC"
FT Region 909..1095
FT /label= B3
FT /note="see CC"
FT Region 1096..1159
FT /label= W
FT /note="see CC"
FT Region 1160..1179
FT /label= M
FT /note="see CC"
FT Region 1180..1185
FT /note="see CC"
XX /note="charged C-terminal"
XX MO9207002-A.

```

```

XX PD 30-APR-1992.
XX PF 22-OCT-1991; 91MO-SE00207.
XX PR 22-OCT-1990; 90SE-0003374.
XX PA (ALFA ) ALFA LAVAL AGRI INT AB.
XX PI Guss BM, Hook M, Jonsson H, Lindberg KM, Patti J, Signaes LC;
XX PI Switalski LM;
XX DR MPI: 1992-167099/20.
XX DR N-PSDB; Q24123.
XX PT Hybrid DNA molecule encoding S.aureus collagen binding protein -
XX PT protein is expressed in E.coli and used for diagnosis e.g. of
XX PT septic arthritis
XX PS Disclosure; Fig 2; 40pp; English.
XX CC The amino acid sequence (encoded by the sequence assembled from
XX CC partially homologous p16 and cCOLR6A inserts) comprises a structure
XX CC resembling a signal sequence. Following this region, a region
XX CC called A is found followed by a repetitive stretch of 187 amino
XX CC acids called B1, B2 and B3. Directly following these regions there
XX CC is a region called W which consists of a repetitive, hydrophilic
XX CC structure contg. several proline residues. This region is thought
XX CC to mediate the binding of the protein to the cell wall. The amino
XX CC acid sequence nearest to the C-terminal end consists of a long
XX CC stretch of hydrophobic residues followed by some charged amino acids
XX CC This region is called M.
XX CC The CBP can be used for immunisation pref. in combination with a
XX CC fusion protein, e.g. for vaccination of ruminants against mastitis
XX CC caused by staphylococcal infections. It can also be used to block
XX CC infection in an open skin wound, e.g. for blocking protein receptors
XX CC or by immunisation. In the latter, the host produces specific
XX CC antibodies which block the adherence of the bacterial strains to
XX CC damaged tissue. This treatment can be used for septic arthritis
XX CC and tissue damage of e.g. skin, connective tissue, and mucous
XX CC membranes. Dosage for immunisation is 0.5-5 microg CBP/kg; for
XX CC topical admin. the protein is used at a concn. of 25-250 microg/ml.
XX SQ Sequence 1185 AA;
QY 98 LEGSRSY-----QVCFENLK-----KAPPLGSDSSVKK-W----- 126
Db 395 leapryftldakkeypfmldtngyftlienakalektkdsaqkvegqtkvpxtly 454
QY 127 ---YKKHDGIST-----KREDYAMSPRITGDDL---NOKLRVAVN----- 160
Db 455 fklxqddngntprvdkaeklkled--gltkvwslnpndkngkaikylvkevnagged 512
QY 161 ---NGHPQANANGIM--EGLEPLNAIRVTOEAVVYVYEDNAPISNPDSEFRSESNLVSTS 215
Db 513 ttpgytltkenglvntentekpietstisgskvwddkndngdkrpek-----vsyn 562
QY 216 QLSLMRQALKQLIDNLAIRKMPKQVPDDFQLSIFESBDGDK-----YNGGYONLLS 267
Db 563 llangekv--ktldvsetnwkyefkd-----lpkydegkkleyvtbedhvytdtldn 614
QY 268 GGLVPTKPPPPG-----DPPMPPNQPPQTTVLIRKAYAGDSKLLKSGATLDLQTD 317
Db 615 gtli-tnkytpgelsatvknwddnnngdkripetelkvelydgqkat---qktalines 669
QY 318 NVNSFOARVFSSNDIGERIELSDGTYTLTELNSPAGYSI-----ABPIPF 362
Db 670 nwthtwtgldekakgqykv-----ytveelkvkygtytchvdnmgnlvtnkypelt 724

```

```

QY 363 KVEAGVYTIIDGKQIENPKELIVEPYSEAYNDPEFSYLTQNAKEYY-----AKNK 417
DB 725 slsgekvwdkndqgdkrpk-----vsynlladgkvktlvtsetnwkyefkdlpkypd 779
QY 418 NGSSQVVCYFNADLKSPPESEDCGKTMTPTFTGVEVKYTHIAGRD-----462
DB 780 eg-kkleyctvehvdyttdingttltnkytgsatsatvkwnndnngdqrrpeikv 838
QY 463 -LF---KYVKKPDTDPDTFLKIKVIEKGYREKGOALYSGLTQLRAAQ-----512
DB 839 elyqdgkatsktaltnesnmwhtwtgide--kakgqvkytveeltkvkgylthvdnd 896
QY 513 -----LAIFYTSAELDKDKLDYHGFGDMNSTLAVAKILVEYADSNPQOLDLDF 567
DB 897 mgnlitvtnkypetstaisgekwad-----dkan-----qdgkrpekvsn-1 937
QY 568 IPNNKRYOSL---IGTQWPEDELVDIIRMEDKKEVIVPHTNLRTKVTGLADRRKDFH 634
DB 938 langekvktldvsetnwkye-fkdlpkysdegkkl-----eytvt---edhvkdyt 984
QY 625 FELE-----LKNKKQELISQVKT---DKTNLE-----FKDGKAT-----INLKH 661
DB 985 tdingttltnkytgsatsatvkwnndnngdqrrpeikvlyqdgkatsktaltnesn 1044
QY 662 GSELTQGLPE-----GYSYLVE-TDSEGYKVKVNSOEYANATVSK--TGINSDELAF 713
DB 1045 nwthtwtgldkakgqgvkytvdeltkngylthvdnndmgnlitvtnkypkpknpkyp 1104
QY 714 ENNKEPVPTGVD 726
DB 1105 ekpkdktpkpkp 1117

RESULT 15
Y08603 ID Y08603 standard; Protein: 1112 AA.
XX AC Y08603;
XX DT 05-AUG-1999 (first entry)
XX DE
XX DE S. pyogenes SFEBP-12 protein.
XX KM SFEBP-12: fibronectin: fibrinogen: group A Streptococci: infection:
XX KM fibrinogen and fibronectin binding protein: bacterial adhesion: vaccine:
XX KM diagnosis: treatment: prevention: streptococcal infection: antigen:
XX KM immune system: etiologic agent: suppurative infection: pharyngitis;
XX KM impetigo: necrotizing fasciitis; systemic disease: scarlet fever;
XX KM toxic-shock syndrome; sequelae; rheumatic fever; glomerulonephritis.
XX OS Streptococcus pyogenes.
XX PN US5910441-A.
XX PD 08-JUN-1999.
XX PF 16-SEP-1996; 96US-0714402.
XX PR 16-SEP-1996; 96US-0714402.
XX PA (UYRQ ) UNIV ROCKEFELLER.
XX PI Fischetti VA, Rocha C;
XX DR WPI; 1999-356822/30.
XX DR N-PSDB; X77451.
XX PT Isolated DNA that encodes group A Streptococci fibrinogen and
XX PS and treatment of Streptococcal infections
Claim 1; Fig 4; 17pp; English.

```

```

XX This invention describes the isolation of a novel group A Streptococci
CC fibrinogen and fibronectin binding protein (SFEBP-12) and its encoding
CC nucleic acid which is involved in adhesion of the bacteria to the host
CC cell, and may initiate the infection process). The products of the
CC invention are useful in the preparation of compositions and tests for
CC the diagnosis, treatment and prevention of streptococcal infection.
CC SFEBP-12 nucleic acid may be used as a hybridization probe to isolate
CC corresponding genes from other species or in biological samples by, for
CC example, Northern/southern blotting. Additionally, the nucleic acid may
CC also be transfected into host cells and used to recombinantly produce
CC SFEBP-12 proteins in fermentation cultures. The SFEBP-12 protein may be
CC used as an antigen in the preparation of vaccines to stimulate a hosts
CC immune system against Streptococcal infection. Group A Streptococci (e.g.
CC Streptococcus pyogenes) are the etiologic agents for a range of
CC suppurative infections (e.g. pharyngitis, impetigo and necrotizing
CC fasciitis), systemic diseases (e.g. scarlet fever, toxic-shock syndrome)
CC and may lead to serious sequelae (e.g. rheumatic fever and
CC glomerulonephritis).
XX
SQ Sequence 1112 AA;

```

```

Query Match 3.7%; Score 145; DB 20; Length 1112;
Best Local Similarity 19.8%; Pred. No. 0.063;
Matches 190; Conservative 121; Mismatches 327; Indels 324; Gaps 51;

```

```

QY 21 KNSKRTYVLY--GVFLMT-----FALVTSNVCAGKTYFGLVESTPRAINPDSSEYRW--- 72
DB 112 kstvtwtvtyengytkllyvnpngelislkagskdv-----ssslqlenpkmsvskyyk 166
QY 73 ---YGESYVRGHPYR---KQ-----FRVADHLRVNLEGSRSYGVYCF 109
DB 167 tewsgaadtyrnhaayfkmfslkqdksetlmpdgtlvtlqdrlnpkq-----217
QY 110 NLKRAP-----LGSDDSVKWKYKRDGISTKEDYA-----141
DB 218 -lsqdlprklllydeansprlaigkyhaenhqlylftdyagldkvqsaelslflenkev1 276
QY 142 -----MSPRINGDELNCK--LRAYMYNGHPONANGIMEGL-----EPNAIVTQE 185
DB 277 entslnfstlsggqelkygvtvnlygnestkesytlnglsnvgysiesynt--etge 334
QY 186 AWVYSDNAPISN-PDESEK-----RESENIWVTSQSLMRQALKOLIDPNATKMP 237
DB 335 fvyvynvnpnrtnlpyatmlmgfgrarntsldendantsselselqyvepegek1p 394
QY 238 KOVPDD-----FQSLFESSEKDKK--YNGK-----262
DB 395 ssygvvtlkltrtdltaglngfmgtkrgldfmglnqkalfllkvgtktdsgskplv 454
QY 263 -ONLLS-----GG-----LVPTKRPPTGDDPMPNPNQOTSVLIRKYA 299
DB 455 qsnlasfrgaseyaftpyvggnvyfqnelaipsksgsgsksef--tkpslvanikrva 512
QY 300 IGDYSKL-----LEGATQLTGDNVNSFQARFSSNDIGE--RIELSDGYTLTELNSP 351
DB 513 qlrfkkmstdnvp1paaatelsngns-qkleassntgvevfkltsgetydllyetkcp 571
QY 352 AGYS-----IAEPITF-----KYBAGKVYTIIDGKO-IENPKELIVE 387
DB 572 kygyvteklatvtvdtkpaemvltwsgpsvskveannevltivhketltsfgk1we 631
QY 388 PYSEAYNDPEFSYLTQNAKEYFAKKNKNGSSQVVCYFNADLKSPPESEDCGKTMTPD 447
DB 632 -----ndrpd-----qpakiqvqllqngq-----kmpnq1ge-----vtkd 663
QY 448 FTTGVEVKYTHIAGRDLFKYVKKPDTDPDTFLKIKVIEKGYREKGOALYSGLTQLRAAQ 507
DB 664 ---ndwsy-hf--kdlpkydakng-----ykyvseevnvpqgkvsyngndifnt 708
QY 508 RAATQALAIYFTDSAEIDKDKLDYHGFGDMNSTLAVAKILVEYADSNPQOLDLDF 565

```

Db 709 r-eteifveqnfnefngnaelkqsgskl1deedltscfkgk1wkn0taenprqalqvq 767
QY 566 FE-----IPNNNKYQSLIGTQWHE-----DLV-----D 589
Db 768 lyadvavegqtkfiagsgnewsefknllkkyngtgn0lysvkevltvplvgdvty sand 827
QY 590 IIRMEDKKEVI-----PYTHNLTLRKTYTGLAGDRTKDFHEIE---LKNNKQELLS 638
Db 828 11--ntkreveltqgpkleleelpllesgasg--gttivedsrpvdclsglisseqgsgd 883
QY 639 QTVKTDK-TNLEFK---DGK---ATINLKHGESLTLOG-----LPEGKSYL 677
Db 884 mtleedsaathikfekrdidgkelagatmelrdsqgkltstvwisdgqvkdflympgkytf- 942
QY 678 VKETDSEGYV-----KVSQEVANATVSKTGITSDETLAFENNKEPVVPFG---VDQ 727
Db 943 vetaapdgyelata1ftvneq--gyvtvngkaekgdtlhwvdaykpkkysgqvdlde 1000
QY 728 KI 729
Db 1001 KI 1002

Search completed: June 7, 2001, 00:17:04
Job time: 2962 sec

